Assignment 18: DIY Least Square method for linear data fitting

You have measured a set of data points, $\{x_i, y_i\}, i = 1,2, \ldots, N$; and you know that they should approximately lie on a straight line of the form *y* = *a* x + *b* if the y_i's are plotted against x_i's. We wish to know what are the best values for *a* and *b* that make the best fit for the data set. The process is called 'data fitting'. The function to be fit against is in a linear form, $y = a + bx$.

Least square method provides a way to obtain the best fit values for the parameters a and b, and their corresponding standard errors:

$$a = \frac{\bar{y}\left(\sum_{i=1}^{n} x_i^2\right) - \bar{x}\sum_{i=1}^{n} x_i\, y_i}{\sum_{i=1}^{n} x_i^2 - n\,\bar{x}^2}$$

$$b = \frac{\left(\sum_{i=1}^{n} x_i\, y_i\right) - n\,\bar{x}\,\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\,\bar{x}^2}$$

$$\bar{x} = \frac{1}{N}\sum_{i}^{N} x_i\,, \bar{y} = \frac{1}{N}\sum_{i}^{N} y_i$$

$$\mathrm{ss}_{x\,x} = \sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$\mathrm{ss}_{x\,y} = \sum_{i=1}^{n}(x_i - \bar{x})\,(y_i - \bar{y})$$

$$\mathrm{ss}_{y\,y} = \sum_{i=1}^{n}(y_i - \bar{y})^2$$

$$s = \sqrt{\frac{\mathrm{ss}_{y\,y} - b\,\mathrm{ss}_{x\,y}}{n-2}} = \sqrt{\frac{\mathrm{ss}_{y\,y} - \frac{\mathrm{ss}_{x\,y}^2}{\mathrm{ss}_{x\,x}}}{n-2}}$$

$$\mathrm{SE}(a) = s\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\mathrm{ss}_{x\,x}}} \qquad \mathrm{SE}(b) = \frac{s}{\sqrt{\mathrm{ss}_{x\,x}}}.$$

Standard errors in a and b are given by $\mathrm{SE}(a)$ and $\mathrm{SE}(b)$

Q1:

- Download the data "data_for_linear_fit.dat" online.

- Develop a Mathematica code that will automatically import the data set to calculate the best-fit values of a and b, and their corresponding standard errors, based on the least-square method formula shown in previous page.

- You should also show in your code the results $(a, b, \mathrm{SE}(a), \mathrm{SE}(b))$ obtained by using the Mathematica built-in function $\mathbf{NonlinearModelFit}[\ ]$ or $\mathbf{LinearModelFit}[\ ]$. Both the results obtained from your DIY code and that from the built-in function should agree to each other.